

## ANALISIS BUTIR SOAL UJIAN SEMESTER GENAP MATA PELAJARAN KIMIA KELAS X IPA SMA PGRI PEKANBARU TAHUN AJARAN 2013/2014

Vemi Yonelia<sup>\*</sup>, Sri Haryati<sup>\*\*</sup> dan Johni Azmi<sup>\*\*\*</sup>

<sup>\*</sup>email: [vmy.254@gmail.com](mailto:vmy.254@gmail.com), Hp: 082392956979

Program Studi Pendidikan Kimia  
Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau

**Abstract:** *Evaluation is one of the component parts that holds a very important role in learning system. One evaluation conducted is evaluation students learning result. The evaluation students learning result generally ranked students on normal curve, so that most of the students are or approaching the median (KKM), as well as a small portion is above the average and below the average. Information was obtained from high school PGRI Pekanbaru, even semester examination results grade X IPA on chemical subjects classified was not ranked students on a normal curve. There are 69 students who reach scores of above KKM and only 1 students under KKM. The KKM score is 80 for students to get completeness . Until in fact the exam result curve was obtained a curve right set. So it needs to be done analysis grains of question which has been used on that examination. Analysis grains of question was done as qualitative and quantitative analysis. The result of the qualitative analysis of the 40 questions, there was only 10 questions (25% ) satisfied aspect of the material, the aspect of the construction and language. The propotion of the division on each chapter not balanced and the distribution of the domain of cognitive taxonomic Bloom has not spread equally. Results of the quantitative analysis of 40 questions, there was 1 questions (2,5%) with a category acceptable, 15 questions (37,5%) with categories is fixed and 24 questions (60%) were rejected by category.*

**Key words:** *analysis grain of questions, quality, quantity*

## **ANALISIS BUTIR SOAL UJIAN SEMESTER GENAP MATA PELAJARAN KIMIA KELAS X IPA SMA PGRI PEKANBARU TAHUN AJARAN 2013/2014**

Vemi Yonelia<sup>\*</sup>, Sri Haryati<sup>\*\*</sup> dan Johni Azmi<sup>\*\*\*</sup>

<sup>\*</sup>email: [vmy.254@gmail.com](mailto:vmy.254@gmail.com), Hp: 082392956979

Program Studi Pendidikan Kimia  
Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau

**Abstrak:** Evaluasi merupakan salah satu komponen yang memegang peranan yang sangat penting dalam sistem pembelajaran. Salah satu evaluasi yang dilakukan adalah evaluasi hasil belajar siswa. Hasil evaluasi belajar siswa pada umumnya memposisikan siswa pada kurva normal, sehingga sebagian besar siswa berada atau mendekati garis rata-rata (KKM), serta sebagian kecil berada di atas rata-rata dan di bawah rata-rata. Data yang diperoleh dari SMA PGRI Pekanbaru menunjukkan bahwa hasil ujian semester genap mata pelajaran kimia kelas X IPA tidak memposisikan siswa pada kurva normal. Terdapat 69 siswa yang memperoleh nilai di atas KKM dan 1 orang siswa yang memperoleh nilai di bawah KKM. nilai KKM yang harus dicapai siswa untuk mencapai ketuntasan adalah 80. Sehingga kurva hasil evaluasi belajar siswa yang diperoleh berbentuk kurva juring ke kanan. Oleh karena itu perlu dilakukan analisis terhadap butir-butir soal yang digunakan pada ujian tersebut. Analisis butir soal dilakukan secara kualitatif dan kuantitatif. Hasil analisis secara kualitatif dari 40 soal diperoleh 10 soal (25%) yang telah memenuhi kriteria aspek materi, konstruksi dan bahasa. Proporsi pembagian soal pada setiap pokok bahasan tidak seimbang dan distribusi jenjang ranah kognitif taksonomi Bloom belum tersebar merata. Hasil analisis kuantitatif dari 40 soal diperoleh 1 soal (2,5%) dengan kriteria diterima, 15 soal (37,5%) dengan kriteria diperbaiki dan 24 soal (60%) dengan kriteria ditolak.

**Kata Kunci:** Analisis butir soal, kualitatif, kuantitatif

## PENDAHULUAN

Evaluasi merupakan salah satu komponen yang memegang peranan yang sangat penting dalam sistem pembelajaran. Hal ini disebabkan karena, dengan melakukan evaluasi keberhasilan proses pengajaran dapat diketahui. Salah satu evaluasi yang dilakukan adalah evaluasi hasil belajar siswa. Evaluasi hasil belajar siswa dapat menggambarkan kemampuan dan tingkat keberhasilan siswa dalam pencapaian tujuan pembelajaran, setelah siswa menempuh proses pembelajaran dalam jangka waktu yang telah ditetapkan (Anas Sudijono, 2011).

Evaluasi hasil belajar siswa dilakukan dalam bentuk ujian harian, ujian tengah semester, ujian akhir semester dan ujian kenaikan kelas. Salah satu instrumen yang digunakan sebagai sarana untuk evaluasi hasil belajar adalah tes. Saifuddin Azwar (2002) menyatakan bahwa, “Tes adalah sekumpulan pertanyaan yang dapat mengungkap keberhasilan seseorang dalam belajar”. Tes pada umumnya digunakan untuk menilai dan mengukur hasil belajar siswa, terutama hasil belajar kognitif berkenaan dengan penguasaan bahan pengajaran atau ketuntasan siswa terhadap pelajaran yang telah diajarkan.

Hasil wawancara dengan guru pengampu mata pelajaran kimia kelas X IPA di SMA PGRI Pekanbaru, dapat diketahui bahwa hasil ujian semester genap mata pelajaran kimia kelas X IPA sangat memuaskan. Hanya satu orang dari 70 siswa yang memperoleh nilai di bawah Kriteria Ketuntasan Minimum (KKM) yaitu 80. E. Mulyasa (2014) menyatakan bahwa “Hasil penilaian ketuntasan belajar siswa pada umumnya memposisikan peserta didik pada kurva normal, sehingga sebagian besar peserta didik berada atau mendekati garis rata-rata (KKM), serta sebagian kecil berada di atas rata-rata dan di bawah rata-rata”. Hasil ujian siswa pada mata pelajaran kimia kelas X IPA menunjukkan bahwa sebagian besar siswa berada di atas garis rata-rata (KKM), hanya satu orang saja yang nilainya berada di bawah rata-rata. Hal ini bertentangan dengan teori yang dikemukakan oleh E. Mulyasa dalam buku Pengembangan dan Implementasi Kurikulum 2013. Permasalahan ini menarik perhatian penulis untuk meneliti lebih lanjut tentang fenomena tersebut. Hal yang sangat erat kaitannya dengan hasil ujian adalah perangkat soal yang digunakan.

Soal ujian semester genap mata pelajaran kimia tahun ajaran 2013/2014 yang dilaksanakan di SMA PGRI Pekanbaru berupa 40 buah soal objektif (pilihan ganda) yang disusun sendiri oleh guru pengampu bidang studi kimia di sekolah tersebut. Soal tersebut disusun sesuai dengan tujuan pembelajaran dan deskripsi bahan ajar yang telah diajarkan. Soal yang digunakan sebagai alat evaluasi haruslah memiliki kualitas yang baik. Jika soal yang digunakan mutunya kurang baik, maka akan mengakibatkan terjadinya kesalahan pengukuran terhadap kemampuan siswa (Thoha, 2003). Hamalik (2003) menyatakan bahwa “Salah satu kunci untuk mendapatkan tes hasil belajar yang baik adalah melalui proses kegiatan analisis butir soal. Sesuai dengan langkah-langkah pelaksanaan ujian yang dijelaskan dalam Peraturan Menteri No. 66 Tahun 2013 bahwa “Langkah kedua yang harus dilakukan guru dalam penilaian hasil belajar siswa adalah mengembangkan (menulis, menelaah (menganalisis), dan merevisi) instrumen yang digunakan”.

Analisis butir soal tes dilakukan untuk mengetahui apakah butir soal tes hasil belajar tersebut sudah dapat berfungsi sebagai alat ukur yang valid atau belum”. Soal dikatakan mempunyai kualitas yang baik apabila sesuai dengan kompetensi dasar, memenuhi syarat aspek materi, konstruksi dan bahasa, mempunyai validitas, reliabilitas dan daya pembeda yang tinggi, tingkat kesukaran yang sedang serta distraktor yang

berfungsi efektif (Suharsimi Arikunto, 2012). Sehingga dapat diketahui butir-butir soal yang perlu diperbaiki ataupun yang diganti serta soal yang dapat digunakan kembali sebagai tes hasil belajar pada periode berikutnya.

Sesuai dengan latar belakang yang telah diuraikan, maka penulis tertarik untuk melakukan penelitian dengan judul ” Analisis Butir Soal Ujian Semester genap Mata Pelajaran Kimia Kelas X IPA SMA PGRI Pekanbaru Tahun Ajaran 2013/2014”. Dengan dilakukannya penelitian ini diharapkan dapat mengetahui butir-butir soal yang perlu diperbaiki ataupun yang diganti serta soal yang dapat digunakan kembali sebagai tes hasil belajar pada periode berikutnya.

## METODE PENELITIAN

Penelitian telah dilaksanakan pada bulan Mei - Desember di Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau (FKIP UR) Pekanbaru. Pengambilan data dilakukan pada minggu keempat bulan Agustus tahun 2014 di SMA PGRI Pekanbaru. Rancangan penelitian berupa penelitian deskriptif yang bersifat *ex post facto*. Objek penelitian adalah silabus, kisi-kisi soal, lembar soal dan lembar jawaban siswa pada ujian semester genap mata pelajaran kimia kelas X IPA SMA PGRI Pekanbaru tahun ajaran 2013/2014.

Analisis soal ujian semester ganjil mata pelajaran kimia kelas X IPA SMA Muhammadiyah 1 Pekanbaru dilakukan secara kualitatif dan kuantitatif. Analisis secara kualitatif dilihat dari segi materi, konstruksi, bahasa dan distribusi ranah kognitif taksonomi Bloom. Analisis kuantitatif butir soal dilakukan secara statistik yaitu analisis tingkat kesukaran, daya pembeda, efektifitas distraktor, validitas butir soal dan reliabilitas soal.

### 1. Analisis Kualitatif

#### a. Analisis Soal dari aspek Materi, Konstruksi dan Bahasa

Untuk membantu analisis soal dapat dibuat kartu telaah soal atau tabel rangkuman telaah soal seperti yang terlihat pada Tabel 1.

Tabel 1. Analisis Soal dari Aspek Materi, Konstruksi dan Bahasa

Aspek yang di telaah	Kriteria
Materi	<ol style="list-style-type: none"> <li>1. Proporsi pembagian jumlah soal yang seimbang sesuai dengan jumlah jam pelajaran tiap pokok bahasan</li> <li>2. Kesesuaian soal dengan indikator</li> <li>3. Kesesuaian materi yang ditanyakan dengan keterpakaian sehari-hari</li> <li>4. Pilihan jawaban homogen dan logis</li> <li>5. Hanya ada satu kunci jawaban</li> </ol>
Konstruksi	<ol style="list-style-type: none"> <li>1. Pokok soal dirumuskan dengan singkat, jelas dan tegas</li> <li>2. Rumusan pokok soal dan pilihan jawaban merupakan pernyataan yang diperlukan saja</li> <li>3. Pokok soal tidak memberi petunjuk kunci jawaban</li> <li>4. Pokok soal besar dari pernyataan yang bersifat negatif ganda</li> <li>5. Gambar, grafik, tabel, diagram atau sejenisnya berfungsi</li> </ol>

Aspek yang di telaah	Kriteria
Bahasa	6. Panjang pilihan jawaban relatif sama
	7. Pilihan jawaban tidak menggunakan “semua jawaban diatas salah/benar”
	8. Pilihan jawaban yang berbentuk angka/waktu disusun berdasarkan urutan besar kecilnya angka atau kronologisnya
	9. Butir jawaban soal tidak bergantung pada jawaban soal sebelumnya
	1. Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia
	2. Menggunakan bahasa yang komunikatif
	3. Tidak menggunakan bahasa yang berlaku setempat/tabu
	4. Pilihan jawaban tidak mengulang kata/kelompok kata yang sama, kecuali merupakan satu kesatuan pengertian

(Silverius, 1991)

Analisis soal dilakukan dengan mencocokkan rumusan soal dengan kriteria yang ada. Apabila rumusan soal sesuai dengan kriteria yang ada maka diberi tanda (✓) pada kolom soal. Jika rumusan soal tidak sesuai dengan kriteria yang ada maka diberi tanda silang (X) pada kolom soal dan jika kriteria yang ada tidak terdapat pada soal maka diberi tanda (-).

b. Analisis distribusi Jenjang Ranah kognitif Taksonomi Bloom

Kualitas butir tes juga dapat dilihat dari tingkat berfikir yang diperlukan dalam mengerjakan soal. Menurut taksonomi Bloom terdapat enam tingkatan ranah kognitif yaitu mengingat (C1), memahami (C2), mengaplikasi (C3), menganalisis (C4), mengevaluasi (C5) dan membuat/mencipta (C6). Analisis distribusi jenjang ranah kognitif taksonomi Bloom dilakukan dengan mencocokkan butir soal dengan kriteria enam tingkatan ranah kognitif taksonomi Bloom.

2. Analisis Kuantitatif

a. Tingkat Kesukaran Soal

Tingkat kesukaran butir soal dihitung dengan menggunakan rumus:

$$P = \frac{B}{JS}$$

Dimana:

P = Tingkat kesukaran

B = Banyak siswa yang menjawab soal itu dengan betul

JS = Jumlah seluruh siswa yang mengikuti tes

(Suharsimi Arikunto, 2012)

b. Daya Pembeda Soal

Daya pembeda butir soal dihitung dengan menggunakan rumus:

$$DP = \frac{B_A}{J_A} - \frac{B_B}{J_B}$$

Dimana:

DP = Daya Pembeda

J = Jumlah peserta tes

- $J_A$  = Banyaknya peserta kelompok atas  
 $J_B$  = Banyaknya peserta kelompok bawah  
 $B_A$  = Banyaknya peserta kelompok atas yang menjawab soal itu dengan benar  
 $B_B$  = Banyaknya peserta kelompok bawah yang menjawab soal itu dengan benar

(Suharsimi Arikunto, 2012)

c. Efektivitas Distraktor

Persentase fungsi butir pengecoh (efektivitas distraktor) dihitung dengan menggunakan rumus:

$$BP = \frac{\text{Jumlah siswa yang memilih butir pengecoh}}{\text{Jumlah siswa peserta tes}} \times 100\%$$

Persentase fungsi butir pengecoh > 5% tergolong baik (Suharsimi Arikunto, 2012).

d. Validitas Butir Soal

Koefisien validitas butir soal dihitung dengan menggunakan rumus:

$$r_{pbi} = \frac{M_p - M_t}{SD_t} \sqrt{\frac{p}{q}}$$

Dimana:

- $r_{pbi}$  = koefisien validitas item  
 $M_p$  = rerata skor dari subjek yang menjawab betul bagi item yang dicari validitasnya  
 $M_t$  = rerata skor total  
 $SD_t$  = Standar deviasi dari skor total  
 $p$  = proporsi siswa yang menjawab benar

$$p = \frac{\text{banyak siswa yang menjawab soal benar}}{\text{jumlah seluruh siswa}}$$

$$q = \text{proporsi siswa yang menjawab salah } q = 1 - p$$

Untuk mengetahui valid atau invalid suatu butir soal dapat dilihat dalam pemberian interpretasi terhadap  $r_{pbi}$ . Dalam pemberian interpretasi terhadap  $r_{pbi}$  digunakan db sebesar (N-2). Derajat kebebasan dikonsultasikan kepada tabel nilai r (product moment), pada taraf signifikan 5%.

- Jika  $r_{pbi} \geq r_{tabel}$ , berarti korelasi tersebut signifikan; artinya butir soal tersebut dikatakan valid
- Jika  $r_{pbi} < r_{tabel}$ , berarti korelasi tersebut tidak signifikan; artinya butir soal tersebut dikatakan invalid

(Anas Sudijono, 2011)

Selanjutnya, butir-butir soal yang telah dianalisis diklasifikasikan berdasarkan kategori “diterima”, “diperbaiki”, “ditolak”.

e. Reliabilitas

Reliabilitas dihitung dengan menggunakan pendekatan konsistensi internal. Formula yang digunakan adalah formula Sperman-Brown yang dikenal dengan istilah teknik belah dua. Model pembelahan yang digunakan adalah model pembelahan awal-akhir. Sebelum menghitung reliabilitas tes, terlebih dahulu

dilakukan korelasi hasil pengukuran kedua kelompok soal menggunakan rumus korelasi *product momen Pearson*, yaitu:

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Koefisien korelasi dengan rumus korelasi *product momen Pearson* akan merupakan korelasi dari setengah jumlah seluruh soal. Untuk mengetahui tingkat reliabilitas keseluruhan soal, dicari koefisien korelasi dengan menggunakan rumus Spearman brown dengan pembelahan awal-akhir yaitu sebagai berikut:

$$r_{11} = \frac{2 \cdot r_s}{1 + r_s}$$

Dimana:

$r_s$  = koefisien antara skor-skor setiap belahan tes

$r_{11}$  = koefisien reliabilitas yang sudah disesuaikan

## HASIL DAN PEMBAHASAN

### A. Hasil

#### 1. Hasil Analisis Kualitatif

##### a. Analisis dari Aspek Materi, Konstruksi dan Bahasa

Hasil Analisis terhadap 40 butir soal diperoleh proporsi pembagian jumlah soal untuk tiap pokok bahasan seperti yang ditunjukkan pada Tabel 2.

Tabel 2. Proporsi Pembagian Jumlah Soal untuk Setiap Pokok Bahasan

No	Pokok Bahasan	Jumlah Jam Pelajaran	Jumlah Soal Seharusnya	Jumlah Soal yang Ada
1.	Larutan elektrolit dan nonelektrolit	4 JP	4	6
2.	Reaksi reduksi – oksidasi	12 JP	11	9
3.	Tata nama senyawa dan persamaan reaksi	6 JP	6	6
4.	Hukum - hukum dasar kimia	8 JP	8	8
5.	Konsep mol dan perhitungan kimia	12 JP	11	11
<b>Total</b>		<b>42 JP</b>	<b>40</b>	<b>40</b>

Hasil analisis kualitatif dari aspek materi, konstruksi dan bahasa dapat dilihat pada Tabel 3.

Tabel 3. Analisis Butir Soal dari Aspek Materi, Konstruksi dan Bahasa

No	Aspek yang Ditelaah	Jumlah Soal yang Tidak Sesuai Kriteria	Jumlah Soal yang dianalisis	Persentase
1.	Aspek Materi	12	40	30 %
2.	Aspek konstruksi	17	40	42,5 %
4.	Aspek Bahasa	17	40	42,5 %

### b. Analisis Distribusi Jenjang Ranah Kognitif Taksonomi Bloom

Analisis distribusi jenjang ranah kognitif Taksonomi Bloom dapat dilihat pada Tabel 4.

Tabel 4. Analisis Distribusi Jenjang Ranah Kognitif Taksonomi Bloom

Jenjang Ranah Kognitif	Jumlah Soal	Persentase
Mengingat (C1)	7	17,5 %
Memahami (C2)	28	70 %
Mengaplikasi (C3)	0	0 %
Menganalisis (C4)	5	12,5 %
Mengevaluasi (C5)	0	0 %
Membuat/ Mencipta (C6)	0	0 %
<b>Total</b>	<b>40</b>	<b>100%</b>

## 2. Hasil Analisis Kuantitatif

Analisis kuantitatif ditinjau dari aspek tingkat kesukaran, daya pembeda, efektivitas distraktor dan validitas butir soal. Setelah dianalisis, 40 butir soal tersebut diklasifikasikan berdasarkan kategori diterima, diperbaiki dan ditolak seperti ditunjukkan pada Tabel 5.

Tabel 5. Klasifikasi butir soal berdasarkan kategori “diterima”, “diperbaiki” dan “ditolak”

Kategori	Jumlah Soal	Persentase
Soal diterima	1	2,5 %
Soal diperbaiki	15	37,5 %
Soal ditolak	24	60 %
<b>Total</b>	<b>40</b>	<b>100 %</b>

## B. Pembahasan

### 1. Analisis Kualitatif

#### a. Analisis dari Aspek Materi, Konstruksi dan Bahasa

##### 1) Analisis dari Aspek Materi

Pada analisis aspek materi perlu diketahui keseimbangan antara proporsi pembagian jumlah soal untuk setiap pokok bahasan dengan jumlah jam pelajaran yang digunakan. Berdasarkan silabus mata pelajaran kimia untuk kurikulum 2013 pokok bahasan kimia kelas X IPA semester genap adalah larutan elektrolit dan nonelektrolit, reaksi reduksi-oksidasi, tata nama senyawa dan persamaan reaksi, hukum-hukum dasar kimia, konsep mol dan perhitungan kimia. Hasil analisis yang telah dilakukan menunjukkan bahwa pembagian jumlah soal untuk setiap pokok bahasan belum semuanya seimbang dengan jumlah jam pelajaran yang digunakan.

Proporsi pembagian jumlah soal untuk pokok bahasan larutan elektrolit dan nonelektrolit dengan pokok bahasan reaksi reduksi-oksidasi belum seimbang. Hal ini akan mengakibatkan penilaian terhadap hasil belajar siswa tidak merata pada setiap pokok bahasan dilihat dari aspek kognitifnya. Pembuatan soal ujian sebaiknya lebih memperhatikan keseimbangan antara proporsi pembagian jumlah soal dengan jam pelajaran yang digunakan. Suharsimi Arikunto (2012) menyatakan bahwa “Penentuan



imbangan bobot untuk masing-masing materi atau pokok bahasan dilakukan berdasarkan luasnya materi atau kepentingannya untuk diukur”. Pokok bahasan dengan materi yang luas, panjang dan rumit harus diajarkan untuk waktu yang cukup lama. Sehingga jumlah soal yang digunakan untuk menilai hasil belajar siswa pada pokok bahasan tersebut seharusnya juga lebih banyak dibandingkan dengan pokok bahasan yang materinya lebih sempit, sedikit dan sederhana.

Hasil belajar siswa diukur berdasarkan kompetensi dasar yang ada pada silabus. Kompetensi dasar tersebut diuraikan menjadi beberapa indikator, dan kemudian indikator diuraikan lagi menjadi beberapa tujuan pembelajaran. Soal evaluasi hasil belajar dibuat berdasarkan tujuan pembelajaran yang ada dalam Rencana Pelaksanaan Pembelajaran (RPP). Hasil belajar siswa dapat terukur dengan baik apabila soal-soal yang digunakan sesuai dengan tujuan pembelajaran, indikator dan kompetensi dasar. Soal-soal yang tidak sesuai dengan tujuan pembelajaran dan indikator harus diganti, karena soal-soal tersebut tidak dapat berfungsi sebagai alat ukur yang relevan untuk mengukur ketercapaian hasil belajar siswa.

Pada soal ujian semester genap mata pelajaran kimia kelas X IPA di SMA PGRI Pekanbaru tahun ajaran 2013/2014, semua soal yang digunakan telah sesuai dengan tujuan pembelajaran yang tercantum dalam Rencana Pelaksanaan Pembelajaran (RPP). Namun, terdapat 6 soal yang tidak sesuai dengan tuntutan indikator, yaitu soal nomor 8, 9, 25, 26, 39 dan 40. Meskipun soal-soal tersebut sesuai dengan tujuan pembelajaran, soal tersebut harus diganti karena tidak memenuhi tuntutan indikator yang digunakan. Contohnya pada soal nomor 8 dan 9. Sebagaimana dijelaskan oleh Slameto (2001) bahwa salah satu dasar pelaksanaan evaluasi adalah kurikulum. Isi evaluasi harus sesuai dengan materi yang diajarkan seperti tercantum dalam kurikulum yang telah ada. Pada kurikulum 2013, kompetensi dasar pada pokok bahasan reaksi reduksi-oksidasi adalah “Menganalisis perkembangan konsep reaksi reduksi-oksidasi...”. Hal yang sama tercantum dalam indikator dimana siswa dituntut untuk dapat menganalisis perkembangan konsep reaksi oksidasi-reduksi. Soal yang digunakan seharusnya soal yang berada pada jenjang ranah kognitif C4 (menganalisis). Namun pada soal nomor 8 dan 9, tipe soalnya berada pada jenjang ranah kognitif C1 (mengingat). Soal-soal seperti ini harus diganti, karena belum dapat mengukur secara optimal ketercapaian pembelajaran oleh siswa.

Selain itu, soal-soal dimana kesesuaian materi yang ditanyakan dengan kompetensi relevansi, kontinuitas, keterpakaian sehari-hari rendah harus diperbaiki. Soal-soal tersebut dapat membuat siswa bingung dengan materi yang diajarkan, salah konsep dan keliru dalam memahami maksud soal tersebut.

Soal yang pilihan jawabannya tidak homogen dan logis akan mengakibatkan siswa dapat lebih mudah dalam menentukan pilihan jawaban, sehingga distraktor pada soal tersebut tidak berfungsi efektif. Sebuah distraktor dapat dikatakan berfungsi dengan baik apabila distraktor tersebut mempunyai daya tarik yang besar bagi peserta tes yang kurang memahami konsep atau kurang menguasai bahan (Suharsimi Arikunto, 2012). Pilihan

jawaban yang tidak homogen dan logis memungkinkan siswa untuk menebak-nebak jawaban dan bisa jadi tebakan jawaban siswa benar. Sehingga siswa yang belum mengerti dengan materi yang ditanyakan pada soal dapat menjawab soal tersebut dengan benar. Oleh karena itu distraktor harus dibuat semirip mungkin dengan kunci jawaban, agar soal tersebut benar-benar dapat membedakan kemampuan siswa dengan tepat.

Pada analisis kualitatif dari aspek materi terdapat 2 soal yang tidak mempunyai kunci jawaban yaitu soal nomor 21 dan 26. Soal ini akan membuat siswa bingung, karena tidak tahu mana yang benar, pernyataan pada soal atau materi yang telah diajarkan guru. Siswa hanya akan memperkirakan jawaban yang tepat, sehingga hasil evaluasi pembelajaran yang diperoleh hanya berdasarkan pada perkiraan dan keberuntungan siswa, bukan berdasarkan kemampuan siswa.

Terdapat juga 3 soal yang kunci jawabannya keliru, yaitu soal nomor 6, 8, dan 34. Soal-soal seperti ini akan membuat pengukuran terhadap hasil belajar siswa menjadi keliru. Siswa yang menjawab dengan benar akan disalahkan oleh guru, karena jawaban siswa berbeda dengan kunci jawaban. Hal ini akan merugikan siswa yang menjawab benar dan menguntungkan siswa yang tidak sengaja menjawab sesuai dengan kunci jawaban guru.

## **2) Analisis dari Aspek Konstruksi**

Hasil analisis yang diperoleh menunjukkan bahwa masih ada soal yang belum memenuhi kriteria aspek konstruksi. Pada analisis aspek konstruksi, pokok soal harus dirumuskan dengan singkat, jelas dan tegas, karena pokok soal yang bertele-tele (mengandung pernyataan yang tidak diperlukan) dan tidak jelas membuat soal menjadi tidak enak untuk dibaca, sulit dipahami dan dapat menimbulkan penafsiran yang berbeda pada setiap siswa, sehingga terjadi kesalahan dalam menjawab soal tersebut.

Soal yang pilihan jawabannya berbentuk angka/waktu harus disusun berdasarkan urutan besar kecilnya angka atau kronologisnya. Penyusunan pilihan jawaban seperti ini bertujuan agar pilihan jawaban pada soal tersusun secara teratur, rapi dan sistematis. Namun pada soal ujian semester genap mata pelajaran kimia kelas X IPA SMA PGRI Pekanbaru tahun ajaran 2013/2014 terdapat soal-soal yang pilihan jawabannya berbentuk angka namun tidak disusun berdasarkan besar kecilnya angka atau kronologisnya.

Pada soal ujian kimia semester genap kelas X IPA SMA PGRI Pekanbaru terdapat 1 buah soal yang butir soalnya bergantung pada jawaban soal sebelumnya, yaitu soal nomor 39 yang jawabannya bergantung pada jawaban soal nomor 27. Jika siswa dapat menjawab soal nomor 27, maka siswa juga akan dapat menjawab soal nomor 39. Soal seperti ini akan mempermudah siswa dalam menjawab kedua soal tersebut. Siswa dapat menjadikan salah satu jawaban soal sebagai tolak ukur untuk menentukan jawaban yang benar pada soal yang lain.

## **3) Analisis dari Aspek Bahasa**

Hasil analisis yang diperoleh menunjukkan bahwa masih ada soal yang belum memenuhi kriteria aspek bahasa. Soal yang tidak menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia akan membuat soal

kurang bisa dipahami oleh siswa. Selain itu penulisan soal yang sesuai dengan kaidah bahasa Indonesia itu penting untuk mengajarkan siswa bagaimana cara berbahasa Indonesia yang baik dan benar, meskipun dalam mata pelajaran kimia (eksak). Sedangkan soal yang tidak menggunakan bahasa yang komunikatif akan membuat siswa bingung dalam membaca soal, bahkan akan ada yang keliru dalam menjawab soal tersebut.

Secara keseluruhan, terdapat 10 atau 25% soal yang memenuhi semua kriteria dari aspek materi, konstruksi dan bahasa, yaitu soal nomor 1, 11, 13, 16, 17, 18, 20, 30, 35 dan 38. Sedangkan 6 soal atau 15% soal harus diganti karena tidak memenuhi tuntutan indikator, yaitu soal nomor 8, 9, 25, 26, 39 dan 40. Serta 24 soal lainnya atau 60% soal harus diperbaiki dari segi kualitatifnya.

#### **b. Analisis Distribusi Jenjang Ranah Kognitif Taksonomi Bloom**

Analisis distribusi jenjang ranah kognitif taksonomi Bloom dilakukan dengan mencocokkan butir soal dengan kriteria 6 tingkatan ranah kognitif yaitu mengingat (C1), memahami (C2), mengaplikasikan (C3), menganalisis (C4), mengevaluasi (C5) dan mencipta (C6). Hasil analisis kualitatif distribusi jenjang ranah kognitif taksonomi Bloom soal ujian kimia semester genap kelas X IPA SMA PGRI Pekanbaru yang diperoleh adalah sebanyak 7 soal atau 17,5% berada pada jenjang ranah kognitif mengetahui (C1), 28 soal atau 70% berada pada jenjang ranah kognitif memahami (C2), dan 5 soal atau 12,5% berada pada jenjang ranah kognitif menganalisis (C4). Tidak ada soal yang berada pada jenjang ranah kognitif mengaplikasi (C3), mengevaluasi (C5) dan mencipta (C6). Soal lebih didominasi oleh jenjang ranah kognitif C2 yaitu sebanyak 70%.

Berdasarkan tabel alokasi persentase setiap aspek untuk berbagai bidang studi yang ditampilkan Suharsimi Arikunto (2012) dalam buku Dasar-dasar Evaluasi pendidikan, diketahui bahwa persentase aspek kognitif untuk bidang studi ilmu pengetahuan alam SMA adalah ingatan (C1) : pemahaman (C2) : aplikasi dan analisis (C3/C4) adalah 40 : 40 : 20. Berdasarkan persentase tersebut jumlah soal seharusnya untuk jenjang ranah kognitif C1 sebanyak 16 soal, C2 sebanyak 16 soal dan C3/C4 sebanyak 8 soal. Dalam pembuatan soal, selain memperhatikan kualitas soal berdasarkan kriteria aspek materi, konstruksi dan bahasa, pembagian jumlah soal untuk setiap jenjang ranah kognitif juga harus diperhatikan. Hal ini disebabkan karena soal yang diujikan harus dapat mengukur tingkat berfikir siswa secara merata dari tingkat berfikir rendah hingga tingkat berfikir tertinggi, sehingga evaluasi hasil belajar yang dilakukan dapat lebih optimal dan dapat mengukur kemampuan siswa pada setiap tingkat berfikirnya.

## **2. Analisis Kuantitatif**

### **a. Tingkat Kesukaran Soal**

Hasil analisis yang diperoleh menunjukkan sebanyak 38 soal atau 95% soal termasuk kriteria soal mudah, 1 soal atau 2,5% soal termasuk kriteria soal sedang, dan 1 soal atau 2,5% soal termasuk kriteria soal sukar. Tingkat kesukaran yang terukur secara kuantitatif merupakan tingkat kesukaran soal berdasarkan tingkat kemampuan siswa. Artinya 38 soal atau 95% soal dapat diselesaikan dengan mudah oleh siswa. Hanya terdapat 1 soal yang dapat

diselesaikan oleh setengah dari jumlah siswa dan 1 soal yang hanya dapat diselesaikan oleh 1 orang siswa.

Hasil analisis tingkat kesukaran yang diperoleh menunjukkan bahwa sebagian besar soal berada pada kriteria soal mudah. Menurut Nana Sudjana dalam Murbangun Nuswowati dkk (2010), sebaiknya sebuah paket soal yang diberikan kepada siswa memiliki keseimbangan antara soal sukar : sedang : mudah dengan perbandingan 3:4:3 atau 2:5:3. Pada soal ujian semester genap mata pelajaran kimia kelas X IPA SMA PGRI diperoleh perbandingan soal mudah : sedang : sukar adalah 1:38:1, sehingga soal-soal tersebut harus diperbaiki.

Kualitas butir soal berdasarkan tingkat kesukaran dapat dipengaruhi oleh kesalahan pada soal, misalnya kesalahan pada perintah, kesalahan pada opsi jawaban, kesalahan pada kunci jawaban, ilustrasi atau bacaan yang sulit dimengerti, atau mungkin juga materi yang belum diajarkan. Tindak lanjut yang dilakukan dalam perbaikan butir soal hendaknya disesuaikan dengan faktor penyebab buruknya kualitas butir soal tersebut. Salah satu solusi yang ditawarkan adalah dilakukannya perbaikan terhadap butir soal secara kualitatif dan perbaikan persentase soal pada setiap jenjang ranah kognitif.

#### **b. Daya Pembeda Soal**

Daya pembeda soal menunjukkan kemampuan suatu soal untuk membedakan antara siswa yang pandai atau berkemampuan tinggi dengan siswa yang berkemampuan rendah. Jika suatu soal dapat dijawab benar oleh siswa yang berkemampuan tinggi dan siswa yang berkemampuan rendah, maka soal tersebut tidak mempunyai daya pembeda yang bagus. Daya pembeda soal akan bernilai negatif jika soal tersebut dapat dijawab benar oleh siswa yang berkemampuan rendah, namun tidak dapat dijawab oleh siswa yang berkemampuan tinggi.

Hasil analisis kuantitatif daya pembeda soal menunjukkan bahwa terdapat 1 soal atau 2,5% soal dengan daya pembeda yang jelek sekali, yaitu soal nomor 31. Terdapat 1 soal atau 2,5% soal dengan daya pembeda yang cukup baik, yaitu soal nomor 10. Serta terdapat 1 soal atau 2,5% soal dengan daya pembeda yang baik sekali, yaitu soal nomor 34. Sisanya 37 soal termasuk soal dengan daya pembeda yang jelek.

#### **c. Efektifitas Distraktor**

Distraktor merupakan jawaban-jawaban yang salah pada soal pilihan ganda. Distraktor bertujuan untuk mengecoh siswa khususnya siswa yang berkemampuan rendah, sehingga siswa tersebut merasa bimbang dan ragu-ragu dalam menjawab soal. Jika siswa tersebut belum memahami materi yang ditanyakan pada soal, bisa jadi siswa akan memilih distraktor sebagai jawaban yang benar. Distraktor dapat dinyatakan telah berfungsi dengan baik apabila dipilih sekurang-kurangnya oleh 5% dari seluruh peserta tes.

Hasil analisis efektivitas distraktor yang telah dilakukan menunjukkan bahwa 5 *option* atau 2,5% *option* dinyatakan telah berfungsi efektif, sedangkan 195 *option* atau 97,5% *option* dinyatakan masih belum berfungsi efektif. Distraktor yang kurang berfungsi dapat disebabkan karena distraktor yang dibuat tidak homogen dan logis, sehingga siswa sudah tahu bahwa *option* tersebut adalah distraktor. Hal ini juga biasa disebabkan karena tingkat kesukaran butir

soal yang rendah, sehingga hampir semua siswa dapat menentukan jawaban yang benar dengan mudah.

#### **d. Validitas Butir Soal**

Validitas butir soal menunjukkan ketepatan mengukur yang dimiliki oleh sebuah soal dalam mengukur apa yang seharusnya diukur. Butir soal dikatakan telah memiliki validitas yang tinggi apabila terdapat kesesuaian antara skor pada butir soal dengan skor totalnya. Dengan kata lain terdapat korelasi positif antara skor butir soal dengan skor totalnya. Teknik korelasi yang tepat digunakan dalam menghitung validitas butir soal adalah teknik korelasi point biserial. Anas Sudijono (2011) menyatakan bahwa “Apabila variabel I merupakan data diskret murni atau data dikotomik, sedangkan variabel II berupa data kontinu, maka teknik korelasi yang tepat untuk digunakan dalam mencari korelasi antara variabel I dengan variabel II itu adalah teknik korelasi point biserial”. Indeks korelasi yang diperoleh diberi lambang  $r_{pbi}$ .

Untuk mengetahui valid atau invalid suatu butir soal, kita harus menginterpretasikan nilai koefisien validitas item. Interpretasi tersebut menggunakan derajat kebebasan (db) sebesar  $N-2$ .

$$db = N - 2 = 70 - 2 = 68$$

Derajat kebebasan yang diperoleh disesuaikan dengan tabel nilai  $r$  (*product moment*) pada taraf signifikan 5% yaitu 0,232 (lampiran 19 halaman 176). Jika harga koefisien validitas item lebih besar dari 0,232 maka butir soal tersebut dikatakan valid. Pada soal nomor 34, harga koefisien validitas adalah 0,43, maka butir soal tersebut dinyatakan valid.

Hasil analisis kuantitatif validitas butir soal yang telah dilakukan menunjukkan bahwa terdapat 3 soal atau 7,5% soal dengan validitas tinggi (valid), yaitu soal nomor 3, 4 dan 7. Terdapat 5 soal atau 12,5% soal dengan validitas cukup (valid), yaitu soal nomor 2, 10, 11, 14 dan 34. Terdapat 8 soal atau 20% soal dengan validitas rendah (valid), yaitu soal nomor 20, 22, 23, 28, 29, 35, 36 dan 37. Terdapat 3 soal dengan validitas rendah (invalid), yaitu soal nomor 5, 9 dan 21. Serta terdapat 21 soal atau 52,5% soal dengan validitas sangat rendah (invalid), yaitu soal nomor 1, 6, 8, 12, 13, 15, 16, 17, 18, 19, 24, 25, 26, 27, 30, 31, 32, 33, 38, 39 dan 40. Hasil analisis validitas butir soal tersebut menunjukkan bahwa 60% butir soal tidak valid. Artinya butir-butir soal tersebut tidak dapat mengukur apa yang seharusnya diukur oleh soal. Butir soal yang tidak valid harus diperbaiki agar kesalahan pengukuran pada tes selanjutnya tidak terjadi lagi.

Selanjutnya soal-soal yang sudah dianalisis tingkat kesukaran, daya pembeda, efektifitas distraktor dan validitasnya diklasifikasikan berdasarkan kriteria “diterima”, “diperbaiki” dan “ditolak”. Hasil yang diperoleh menunjukkan bahwa terdapat 24 atau 60% soal dengan kriteria “ditolak”, 15 soal atau 37,5% soal dengan kriteria “diperbaiki” dan 1 soal atau 2,5% soal dengan kriteria “diterima”. Soal dengan kriteria “ditolak” harus diganti dengan soal yang baru, sedangkan soal dengan kriteria “diperbaiki” harus diperbaiki sesuai dengan kriteria-kriteria pada aspek materi, konstruksi dan bahasa.

#### **e. Reliabilitas Soal**

Reliabilitas soal menunjukkan sejauh mana soal tersebut dapat dipercaya dan seberapa akurat soal tersebut mengukur kemampuan siswa jika

seandainya dilakukan pengukuran ulang. Tinggi rendahnya reliabilitas soal secara empirik ditunjukkan oleh suatu angka yang disebut koefisien reliabilitas. Koefisien reliabilitas dihitung dengan menggunakan pendekatan konsistensi internal. Pendekatan ini digunakan karena cocok dengan soal ujian semester genap yang telah dianalisa. Pengukuran reliabilitas dengan pendekatan ini dilakukan dengan menggunakan satu alat ukur dan pelaksanaan pengukuran dilakukan sebanyak satu kali saja.

Formula yang digunakan untuk menghitung koefisien korelasi adalah formula Sperman-Brown yang dikenal dengan istilah teknik belah dua, dimana model pembelahan yang digunakan adalah model pembelahan awal-akhir. Pertama harus dihitung dahulu koefisien korelasi setengah tes dengan menggunakan rumus korelasi product moment. Koefisien korelasi setengah tes yang diperoleh adalah 0,0813. Tingkat reliabilitas keseluruhan soal dihitung dengan menggunakan rumus Sperman-Brown, diperoleh hasilnya sebesar 0,15 dengan kriteria “Korelasi sangat rendah”. Hasil yang diperoleh menunjukkan bahwa soal tes yang digunakan pada ujian semester genap mata pelajaran kimia kelas X IPA SMA PGRI Pekanbaru tidak dapat dipercaya. Artinya hasil yang diperoleh dari ujian semester genap tersebut tidak menunjukkan kemampuan siswa “yang sebenarnya”. Oleh karena itu soal ujian tersebut harus diperbaiki agar nantinya dapat mengukur kemampuan siswa yang sebenarnya.

## **SIMPULAN DAN REKOMENDASI**

### **A. Simpulan**

Analisis butir soal ujian semester genap mata pelajaran kimia kelas X IPA SMA PGRI Pekanbaru tahun ajaran 2013/2014 dapat disimpulkan bahwa:

1. Kualitas butir soal secara kualitatif dilihat dari aspek materi, konstruksi dan bahasa, terdapat 10 soal atau 25% soal yang sudah memenuhi semua kriteria aspek materi, konstruksi dan bahasa, sedangkan 6 soal atau 15% soal harus diganti dan 24 soal atau 60% soal belum memenuhi kriteria dan harus diperbaiki.
2. Distribusi jenjang ranah kognitif taksonomi Bloom didominasi oleh soal dengan jenjang ranah kognitif C2 sebanyak 70% soal. Soal dengan jenjang ranah kognitif C3, C5 dan C6 tidak ada.
3. Kualitas butir soal secara kuantitatif ditinjau dari validitas, daya pembeda, efektifitas distraktor dan tingkat kesukaran. Terdapat 1 soal atau 2,5% soal yang diterima, 15 soal atau 37,5% soal yang harus diperbaiki dan 24 soal atau 60% soal yang harus diganti. Ditinjau dari segi reliabilitas soal, soal yang digunakan memiliki tingkat reliabilitas yang sangat rendah atau tidak dapat dipercaya dengan indeks 0,15.

### **B. Rekomendasi**

Analisis butir soal ujian semester genap mata pelajaran kimia kelas X IPA SMA PGRI Pekanbaru tahun ajaran 2013/2014 dapat direkomendasikan beberapa hal sebagai berikut:

1. Soal tes yang digunakan sebaiknya disesuaikan dengan kriteria aspek materi, konstruksi dan bahasa agar soal tes tersebut memiliki kualitas yang baik dan mampu mengukur kemampuan siswa dengan tepat. Proporsi pembagian jumlah soal juga harus diseimbangkan dengan jam pelajaran yang digunakan.

2. Soal tes yang digunakan hendaknya memiliki distribusi jenjang ranah kognitif yang tersebar merata, sehingga soal tersebut dapat mengukur kemampuan siswa dari tingkat yang rendah sampai tingkat yang tinggi.
3. Soal yang digunakan hendaknya disesuaikan dengan tingkat berfikir dan kemampuan siswa, agar soal tersebut dapat berfungsi sebagai alat ukur yang valid dan reliabel.
4. Dalam penyusunan soal, sebaiknya dimantapkan dengan melakukan uji coba soal tes yang akan diujikan, melakukan analisis butir soal dan melakukan pengembangan terhadap tes yang disusun berdasarkan analisis butir soal yang telah dilakukan.

## DAFTAR PUSTAKA

- Anas Sudijono. 2011. *Pengantar Evaluasi Pendidikan*. Raja Grafindo Persada. Jakarta.
- Ari Widodo. 2006. *Revisi Taksonomi Bloom dan Pengembangan Butir Soal*. Puspendik. Jakarta.
- Daryanto. 2012. *Evaluasi Pendidikan*. Rineka Cipta. Jakarta.
- Depdiknas. 2009. *Analisis Butir Soal*. Departemen Pendidikan Nasional. Jakarta.
- E. Mulyasa. 2014. *Pengembangan dan Implementasi Kurikulum 2013*. PT. Remaja Rosdakarya. Bandung.
- Eri Djanuarsih. 2010. Validitas dan Reliabilitas Butir Soal. *E-Jurnal Dinas Pendidikan Kota Surabaya*. 1(1). P.1-12. (Online). <http://dispendik.surabaya.go.id/surabayabelajar/jurnal/199/jurnal-9.pdf>. (diakses 4 Juli 2014).
- Murbangun Nuswowati, Achmad Binadja, Soeprodjo dan Khida Efti Nely Ifada . 2010. Pengaruh Validitas dan Reliabilitas Butir Soal Ulangan Akhir Semester Bidang Studi Kimia Terhadap Pencapaian Kompetensi. *Jurnal Inovasi Pendidikan Kimia*, 4(1). P. 566-573. (Online).<http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CB0QfjAA&url=http%3A%2F%2Fjournal.unnes.ac.id%2Fnju%2Findex.php%2FJIPK%2Farticle%2Fdownload%2F1314%2F1390&ei=naN1VISnJdDiuQSU3oDQDA&usg=AFQjCNGUutarzDKQfnetBNW5Ya EUUYKng&bvm=bv.80642063,d.c2E>. (diakses 20 November 2014).
- Purwanto. 2008. *Evaluasi Hasil Belajar*. Pustaka Pelajar. Surakarta.
- Riduwan. 2009. *Belajar Mudah Penelitian Untuk Guru-Karyawan dan Peneliti Pemula*. Alfabeta. Bandung.
- Saifuddin Azwar. 2002. *Tes Prestasi, Fungsi dan Pengembangan Pengukuran Prestasi*. Pustaka Pelajar. Yogyakarta.
- Silverius, Suke. 1991. *Evaluasi Hasil Belajar dan Umpan balik*. PT. Grasindo. Jakarta.
- Siskha Sofiana. 2010. *Analisis Butir Soal Ujian Kenaikan Kelas Mata Pelajaran Kimia Kelas X SMA Negeri 8 Surakarta Tahun Ajaran 2009/2010*. <http://lib.unnes.ac.id/18752/1/4401409006.pdf>. (diakses 18 April 2014).
- Slameto. 2001. *Evaluasi Pendidikan*. Bumi Aksara. Jakarta.
- Suharsimi Arikunto. 2012. *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara. Jakarta.
- Thoha, Chabib. 2003. *Teknik Evaluasi Pendidikan*. Rineka Cipta. Jakarta.
- Yaya Jakaria. 2009. *Uji Coba Model (Validasi)*. Pusat Penelitian Kebijakan dan Inovasi Pendidikan, Badan Penelitian dan Pengembangan, Depdiknas. Jakarta.
- Zainal Arifin. 2009. *Evaluasi Pembelajaran*. PT. Remaja Rosdakarya. Bandung.